

Regularity Properties of Non-Negative Sparsity Sets

Matthew K. Tam*

July 21, 2016

Abstract

This paper investigates regularity properties of two non-negative sparsity sets: non-negative sparse vectors, and low-rank positive semi-definite matrices. Novel formulae for their Mordukhovich normal cones are given and used to formulate sufficient conditions for non-convex notions of regularity to hold. Our results provide a useful tool for justifying the application of projection methods to certain rank constrained feasibility problems.

1 Introduction

The solutions of many optimization and reconstruction problems admit characterizations in terms of certain sparse objects. For example, it is sometimes possible to uniquely solve under-determined linear systems under addition assumptions of sparsity [12]. The difficulty arising in such formulations is in dealing with poorly behaved *sparsity functionals*. Two important examples of such functionals are the ℓ_0 -“norm” for vectors, and the rank function for matrices. It is well known that sparsity functionals lead to problems involving non-convexity and NP-hard complexity (see, for example, [17, 28]).

A popular approach to addressing the aforementioned difficulty is to employ *convex relaxations* [8, 10, 33], thus allowing for application of industrial strength non-linear solvers. For instance, the ℓ_1 -norm promotes sparsity and has consequently been used as a surrogate for its ℓ_0 counterpart. Such relaxations come with varying strengths and theoretical guarantees. For an introduction to the topic, we refer the reader to [17, Ch. 4]. Whilst one may be able to exactly solve a relaxation, it is not always the case that this translates into a satisfactory sparse solution of the original problem.

An alternative approach involves attempting to deal with the original problem’s non-convexity directly [1, 5, 19], and thus avoiding the potential complication of recovering a sparse solution from a convex relaxation. Here one can typically only give theoretical guarantees which apply locally (*i.e.*, within some neighbourhood of a solution). In these cases, *regularity properties* of the constraint sets play an important role and can often be usefully formulated in the language of *normal cones*.

This paper investigates regularity properties of sparsity sets having additional non-negativity constraints. We focus on two such sets: non-negative sparse vectors, and

*Institut für Numerische und Angewandte Mathematik, Universität Göttingen, 37083 Göttingen, Germany. E-mail: m.tam@math.uni-goettingen.de

low-rank positive semi-definite matrices. Simple, novel formulae for their Mordukhovich normal cones are given, and then used to formulate sufficient conditions to ensure various regularity properties hold. Implications for algorithms and applications are discussed, with particular attention given to *low-rank Euclidean distance matrix reconstruction*.

The remainder of this paper is organized as follows. In Section 2, we introduce notation and recall results which will be of use. In Section 3, we consider the non-negative sparse vector settings, before *lifting* results to their positive semi-definite counterparts. In Section 4, we deduce consequences of the results from the previous two sections including regularity properties for problems having non-negative sparsity sets. Finally, in Section 5, various example applications of problems in which non-negative sparsity sets arise are given.

2 Preliminaries and Notation

Let \mathbb{E} denote a finite dimensional real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$. Throughout this paper we focus on two such spaces. The first is \mathbb{R}^m equipped with the standard inner product. The second is the set of real symmetric $m \times m$ matrices denoted \mathbb{S}^m equipped with inner product

$$\langle X, Y \rangle := \text{tr}(X^\top Y),$$

where $\text{tr}(\cdot)$ (resp. $(\cdot)^\top$) denote the trace (resp. transpose) of matrix. The induced norm is the *Frobenius norm* which is given by

$$\|X\| = \sqrt{\sum_{i=1}^m \sum_{j=1}^m X_{ij}^2}.$$

One may, of course, think of the Frobenius norm as treating matrices as “long vectors”.

The set of *positive* (resp. *negative*) *semi-definite* $m \times m$ matrices is denoted \mathbb{S}_+^m (resp. \mathbb{S}_-^m) and we write $x \succeq 0$ (resp. $x \preceq 0$) to mean $x \in \mathbb{S}_+^m$ (resp. $x \in \mathbb{S}_-^m$). The set of $m \times m$ *orthogonal* (resp. *permutation*) matrices is denoted \mathbb{O}^m (resp. \mathbb{P}^m).

The *projection* mapping onto the set $\Omega \subseteq \mathbb{E}$ is the set-valued mapping $P_\Omega : \mathbb{E} \rightrightarrows \Omega$ given by

$$P_\Omega(x) := \left\{ y \in \Omega : \|x - y\| \leq \inf_{z \in \Omega} \|x - z\| \right\}.$$

When $P_\Omega(x) = \{y\}$ (i.e., $P_\Omega(x)$ is a singleton) we write $P_\Omega(x) = y$.

In finite dimensions the *Mordukhovich normal cone* to the set $\Omega \subseteq \mathbb{E}$ at a point $\bar{x} \in \Omega$ can be represented as

$$N_\Omega(\bar{x}) = \{y \in \mathbb{E} : \exists (x_n), (y_n) \text{ s.t. } x_n \rightarrow \bar{x}, y_n \rightarrow y, y_n \in \mathbb{R}_+(x_n - P_\Omega(x_n))\}.$$

For closed convex sets this simplifies to the classical *convex normal cone* given by

$$N_\Omega^{\text{conv}}(\bar{x}) := \{y \in \mathbb{E} : \langle y, x - \bar{x} \rangle \leq 0, \forall x \in \Omega\},$$

but still remains useful in non-convex settings [27, Ch. 1]. The *proximal normal cone* to the set $\Omega \subseteq \mathbb{E}$ at a point $\bar{x} \in \Omega$ is given by

$$N_{\Omega}^{\text{prox}}(\bar{x}) := \mathbb{R}_+ (P_{\Omega}^{-1}(\bar{x}) - \bar{x}).$$

A brief summary of relations between the three normal cones is given in the following fact (see, for example, [6, Lem. 2.4]).

Fact 2.1 (Normal cone inclusions). *Let $\Omega \subseteq \mathbb{E}$ be non-empty with $\bar{x} \in \Omega$. Then:*

- (a) $N_{\Omega}^{\text{conv}}(\bar{x}) \subseteq N_{\Omega}^{\text{prox}}(\bar{x})$.
- (b) If Ω is closed, then $N_{\Omega}^{\text{prox}}(\bar{x}) \subseteq N_{\Omega}(\bar{x})$.
- (c) If Ω is closed and convex, then $N_{\Omega}^{\text{conv}}(\bar{x}) = N_{\Omega}^{\text{prox}}(\bar{x}) = N_{\Omega}(\bar{x})$.

Remark 2.2. For the full definition of the Mordukhovich normal cone, valid in any Banach space, see [27, Def. 1.1]. The above definition is an equivalent characterization which holds in the finite dimensional case [27, Th. 1.6]. \diamond

Given $X \in \mathbb{S}^m$ denote by $\lambda_j(X)$ the j th largest eigenvalue of X . In this way,

$$\lambda_1(X) \geq \lambda_2(X) \geq \dots \geq \lambda_m(X).$$

The *eigenvalue map* is the function $\lambda : \mathbb{S}^m \rightarrow \mathbb{R}^m$ which maps a symmetric matrix to the m -dimensional vector of its eigenvalues arranged in non-increasing order. That is,

$$\lambda(X) := (\lambda_1(X), \lambda_2(X), \dots, \lambda_m(X)).$$

The *indicator function* of a set $\Omega \subseteq \mathbb{E}$ is the function $\iota_{\Omega} : \mathbb{E} \rightarrow \mathbb{R} \cup \{+\infty\}$ which takes the value 0 on Ω , and $+\infty$ otherwise. A function $f : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ is called *symmetric* if $f(x) = f(\sigma x)$ for all $x \in \text{dom } f := \{x \in \mathbb{R}^m : f(x) < +\infty\}$ and $\sigma \in \mathbb{P}^m$. A function $F : \mathbb{S}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ is called *spectral* if $F(U^{\top} X U) = F(X)$ for all $X \in \text{dom } F$ and $U \in \mathbb{O}^m$. A subset of \mathbb{R}^m (resp. \mathbb{S}^m) is said to be *symmetric* (resp. *spectral*) if and only if its indicator function is symmetric (resp. spectral function).

Symmetric and spectral functions have a natural one-to-one correspondence. The relationship is given by

$$F(X) = (f \circ \lambda)(X), \quad f(x) = F(\text{diag } x), \quad (1)$$

where $\text{diag}(x)$ denotes the $m \times m$ diagonal matrix whose diagonal entries are given by the entries of the vector x . Consequently, many important properties can be transferred between symmetric and spectral functions [11]. For instance, the following fact shows projections onto spectral sets are easily computed whenever the projection onto the corresponding symmetric set is accessible.

Fact 2.3 (Projections onto spectral sets). *Let $K \subseteq \mathbb{R}^m$ be a symmetric set. For any $X \in \mathbb{S}^m$, the projection of X onto the spectral set $\lambda^{-1}(K)$ is given by*

$$P_{\lambda^{-1}(K)}(X) = \{U^{\top}(\text{diag}(y))U : y \in P_K \lambda(X), U \in \mathbb{O}^m(X)\},$$

where the set $\mathbb{O}^m(X) := \{U \in \mathbb{O}^m : X = U^{\top}(\text{diag } \lambda(X))U\}$.

Proof. We prove only the “ \subseteq ” inclusion. A proof of the other inclusion can be found in [25, Th. 21]. Suppose $Y \in P_{\lambda^{-1}(K)}(X)$ but X and Y do not have a simultaneously spectral decomposition. Let $X = U^\top \text{diag } \lambda(X) U$ be an ordered spectral decomposition of X . Since $Y \in \lambda^{-1}(K)$, the vector $\lambda(Y) \in K$ and hence

$$U^\top (\text{diag } \lambda(Y)) U \in \lambda^{-1}(K).$$

By *Fan’s inequality* [7, Th. 1.2.1] and the orthogonality of U ,

$$\begin{aligned} \|X - Y\| &> \|\lambda(X) - \lambda(Y)\| = \|U^\top (\text{diag } \lambda(X) - \text{diag } \lambda(Y)) U\| \\ &= \|X - U^\top (\text{diag } \lambda(Y)) U\|. \end{aligned}$$

This implies $Y \notin P_{\lambda^{-1}(K)}(X)$ which is a contradiction, and completes the proof. \square

The various *sub-differentials* of symmetric and spectral functions are also closely related. The following fact states the equivalence in the context of normal cones, which is the important setting for purposes of this paper.

Fact 2.4 (Normals to spectral sets). *Let $K \subseteq \mathbb{R}^m$ be a closed, symmetric set. For any $X \in \lambda^{-1}(K)$, the Mordukhovich normal cone to the spectral set $\lambda^{-1}(K)$ is given by*

$$N_{\lambda^{-1}(K)}(X) = \{U^\top (\text{diag}(y)) U : y \in N_K(\lambda(X)), U \in \mathbb{O}^m(X)\}.$$

The corresponding result for the proximal normal cone also holds.

Proof. Since K is closed and symmetric, its indicator function ι_K is lower semi-continuous and symmetric. The result follows from [13, Th. 4.2]. \square

For further details on the interplay between symmetric and spectral functions, the reader is referred to works of Lewis and others [7, 11, 14, 22, 23, 32].

3 Normal Cones

In this section we provide formulae for normal cones to sparsity sets with non-negative constraints. Our approach is to first address the non-negative sparse vector settings, before lifting the results to symmetric matrices.

The symmetric function of central interest in this paper is the ℓ_0 -functional denoted $\|\cdot\|_0 : \mathbb{R}^m \rightarrow \{0, 1, \dots, m\}$, which counts the number of non-zero entries of a vector. The corresponding spectral function is the matrix *rank* function. Equivalence (1) becomes

$$\text{rank} = \|\cdot\|_0 \circ \lambda, \quad \|\cdot\|_0 = (\text{rank}) \circ (\text{diag}).$$

Let $s \in \{0, 1, 2, \dots, m\}$. The set of *non-negative sparse vectors* is denoted

$$\mathcal{K}_s := \{x \in \mathbb{R}_+^m : \|x\|_0 \leq s\}.$$

These can be viewed as the *lower-level sets* of the function from $\mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by $x \mapsto \|x\|_0 + \iota_{\mathbb{R}_+^m}(x)$. The set of *low-rank positive semi-definite matrices* is denoted

$$\mathcal{S}_s := \{X \in \mathbb{S}_+^m : \text{rank}(X) \leq s\}.$$

Similarly, these can be viewed as the lower-level sets of the function from $\mathbb{S}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by $X \mapsto \text{rank}(X) + \iota_{\mathbb{S}_+^m}(X)$.

Remark 3.1 (\mathcal{K}_s and \mathcal{S}_s are closed sets). For all $s \in \{0, 1, 2, \dots, m\}$, the set \mathcal{K}_s is closed as the intersection of the two closed sets \mathbb{R}_+^m and $\{x \in \mathbb{R}^m : \|x\|_0 \leq s\}$. The latter is closed being a lower-level set of the lower semi-continuous function $\|\cdot\|_0$. Similarly, \mathcal{S}_s is closed as the intersection of the closed sets \mathbb{S}_+^m and $\{X \in \mathbb{S}^m : \text{rank } X \leq s\}$. The latter being a lower-level set lower semi-continuous function $\text{rank}(\cdot)$. \diamond

3.1 Non-Negative Sparse Vectors

Given a vector $x \in \mathbb{R}^m$ we denote $\mathbb{I}(x) := \{j \in \{1, 2, \dots, m\} : x_j \neq 0\}$. For convenience, we write $x^+ := P_{\mathbb{R}_+^m}(x) = \max\{0, x\}$ (in the pointwise sense) and $x^- := P_{\mathbb{R}_-^m}(x) = \min\{0, x\}$ (see [2] for further details). The standard basis for \mathbb{R}^m is denoted e_1, e_2, \dots, e_m . The set of *sparse vectors* is denoted

$$\mathcal{A}_s := \{x \in \mathbb{R}^m : \|x\|_0 \leq s\}.$$

The following proposition states, in particular, that the projection onto the set \mathcal{K}_s of a vector is given by a simple thresholding of the vector keeping only its s largest non-negative entries.

Proposition 3.2 (Projection onto \mathcal{K}_s and its inverse). *The following hold.*

(a) $\forall x \in \mathbb{R}^m$ and $\forall y \in P_{\mathcal{K}_s}(x)$, $\mathbb{I}(y) \subseteq \mathbb{I}(x^+)$.

(b) $\forall x \in \mathbb{R}^m$, $P_{\mathcal{K}_s}(x) = P_{\mathcal{K}_s}(x^+)$.

(c) $\forall x \in \mathbb{R}_+^m$, $P_{\mathcal{K}_s}(x) = P_{\mathcal{A}_s}(x)$.

(d) $\forall x \in \mathbb{R}^m$, $P_{\mathcal{K}_s}(x) = P_{\mathcal{A}_s}(x^+)$ and hence

$$P_{\mathcal{K}_s}(x) = \left\{ y \in \mathbb{R}^m : y_j = \begin{cases} x_j^+, & j \in \mathbb{J}, \\ 0, & j \notin \mathbb{J}; \end{cases} \text{ for some } \mathbb{J} \in \mathcal{J}_s(x) \right\},$$

where

$$\mathcal{J}_s(x) := \left\{ \mathbb{J} \subseteq \{1, 2, \dots, m\} : |\mathbb{J}| = s, \min_{j \in \mathbb{J}} x_j^+ \geq \max_{j \notin \mathbb{J}} x_j^+ \right\}.$$

(e) If $y \in \mathcal{K}_s$ and $\|y\|_0 = s$, then

$$P_{\mathcal{K}_s}^{-1}(y) = \left\{ x : y_j = x_j \text{ for all } j \in \mathbb{I}(y), \min_{j \in \mathbb{I}(y)} y_j \geq \max_{j \notin \mathbb{I}(y)} x_j^+ \right\}.$$

(f) If $y \in \mathcal{K}_s$ and $\|y\|_0 < s$ then $P_{\mathcal{K}_s}^{-1}(y) = \{x : x^+ = y\} = P_{\mathbb{R}_+^m}^{-1}(y)$.

Proof. (a) Let $y \in P_{\mathcal{K}_s}(x)$ and suppose there exists an index $j_0 \in \mathbb{I}(y) \setminus \mathbb{I}(x^+)$. Then $y_{j_0} > 0$ and $x_{j_0} \leq 0$. Letting $z := y - y_{j_0}e_{j_0} \in \mathcal{K}_s$ we deduce

$$\|x - y\|^2 = \sum_{j \neq j_0} |x_j - y_j|^2 + |x_{j_0} - y_{j_0}|^2 > \sum_{j \neq j_0} |x_j - y_j|^2 + |x_{j_0} - 0|^2 = \|x - z\|^2,$$

which contradicts the assumption $y \in P_{\mathcal{K}_s}(x)$.

(b) Let $y \in P_{\mathcal{K}_s}(x) \cup P_{\mathcal{K}_s}(x^+)$ be arbitrary. By (a), $\mathbb{I}(y) \subseteq \mathbb{I}(x^+)$, hence $\langle x^+ - y, x^- \rangle = 0$ and

$$\|x - y\|^2 = \|x^+ - y\|^2 + \|x^-\|^2.$$

This implies

$$\arg \min_{y \in \mathcal{K}_s} \|x - y\| = \arg \min_{y \in \mathcal{K}_s} \|x^+ - y\|,$$

from which the result follows.

(c) Since $\mathcal{A}_s \cap \mathcal{K}_s = \mathcal{K}_s$, it suffices to show that any $y \in P_{\mathcal{A}_s}(x)$ is contained in \mathcal{K}_s . To this end, suppose $y \in P_{\mathcal{A}_s}(x) \setminus \mathcal{K}_s$ and let j_0 be an index such that $y_{j_0} < 0$. Letting $z := y - y_{j_0} e_{j_0} \in -\mathcal{K}_s$, since $x \in \mathbb{R}_+^m$ we have

$$\|x - y\|^2 > \sum_{j \neq j_0} |x_j - y_j|^2 + |x_j - 0|^2 = \|x - z\|^2,$$

which contradicts the assumption that $y \in P_{\mathcal{A}_s}(x)$.

(d) Follows from (b), (c) and [5, Prop. 3.6(ii)]. (e) Follows from (d) and [5, Prop. 3.6(v)].

(f) Follows from (d) and [5, Prop. 3.6(vi)]. \square

Given a vector $x \in \mathbb{R}^m$ denote by $[x]$ the vector in \mathbb{R}^m obtained by permuting the entries of x in non-increasing order. Under this notation, we note that $[x]_j$, the j th coordinate of the vector $[x]$, is the j th largest entry in the vector x . For vectors $x, y \in \mathbb{R}^m$, $x \odot y$ denotes the *Hadamard product* given pointwise by $(x \odot y)_j := x_j y_j$ for all $j \in \{1, 2, \dots, m\}$. It is worth noting that for fixed \bar{x} , the set $\{y \in \mathbb{R}^m : \bar{x} \odot y = 0\}$, is simply the perpendicular subspace to the support of \bar{x} .

The following lemma is a kind of (non-convex) analogue to *Moreau's decomposition theorem* [3, Th. 6.29], which applies to convex cones, for the set \mathcal{K}_s .

Lemma 3.3 (Decomposition lemma for \mathcal{K}_s). *Let $x, y, z \in \mathbb{R}^m$ with $x = y + z$. Then $y \in P_{\mathcal{K}_s}x$ if and only if $y \in \mathcal{K}_s$, $y \odot z = 0$ and $[y]_s \geq [z]_1$.*

Proof. Suppose $y \in P_{\mathcal{K}_s}x \subseteq \mathcal{K}_s$. By Proposition 3.2 there is an index set $\mathbb{J}_0 \in \mathcal{J}_s(x)$ such that

$$y_j = \begin{cases} x_j^+, & j \in \mathbb{J}_0, \\ 0, & j \notin \mathbb{J}_0; \end{cases} \quad \text{and} \quad \min_{j \in \mathbb{J}_0} x_j^+ \geq \max_{j \notin \mathbb{J}_0} x_j^+.$$

Thus z is given pointwise by

$$z_j = x_j - y_j = \begin{cases} x_j^-, & j \in \mathbb{J}_0, \\ x_j, & j \notin \mathbb{J}_0. \end{cases}$$

It follows that $[y]_s \geq [z]_1$ since

$$\begin{aligned} [y]_s &= \min_{j \in \mathbb{J}_0} x_j^+ \geq \max_{j \notin \mathbb{J}_0} x_j^+ \\ &\geq \max(\{x_j^+ : j \notin \mathbb{J}_0\} \cup \{x_j^- : j \in \mathbb{J}_0\}) \\ &\geq \max(\{x_j : j \notin \mathbb{J}_0\} \cup \{x_j^- : j \in \mathbb{J}_0\}) = [z]_1. \end{aligned}$$

To show $y \odot z = 0$, observe that $(y \odot z)_j = y_j z_j$ is either $x_j^+ x_j^-$ if $j \in \mathbb{J}_0$, or $0 \cdot x_j$ if $j \notin \mathbb{J}_0$, which is zero in either case.

Conversely, suppose $y \in \mathcal{K}_s$, $y \odot z = 0$ and $[y]_s \geq [z]_1$. Since $x = y + z$ and $y \odot z = 0$, for each $j \in \{1, 2, \dots, m\}$ either $y_j = x_j$ and $z_j = 0$; or $y_j = 0$ and $z_j = x_j$. Since $y \in \mathcal{K}_s$ there is an index set \mathbb{J}_0 with $|\mathbb{J}_0| \leq s$ such that we may express

$$y_j = \begin{cases} x_j^+, & j \in \mathbb{J}_0, \\ 0, & j \notin \mathbb{J}_0; \end{cases}, \quad z_j = \begin{cases} x_j^-, & j \in \mathbb{J}_0, \\ x_j, & j \notin \mathbb{J}_0; \end{cases} \quad (2)$$

noting that $x_j = x_j^+$ and $x_j^- = 0$ for $j \in \mathbb{J}_0$.

To show that $y \in P_{\mathcal{K}_s}(x)$ it suffices to consider the case in which $|\mathbb{J}_0| = s$. For if $|\mathbb{J}_0| < s$ then $[y]_s = 0$, hence

$$0 = [y]_s \geq [z]_1 = \max \left(\{x_j^- : j \in \mathbb{J}_0\} \cup \{x_j : j \notin \mathbb{J}_0\} \right).$$

In particular, $x_j \leq 0$ for $j \notin \mathbb{J}_0$, or equivalently $x_j = x_j^-$ for $j \notin \mathbb{J}_0$. It is therefore possible to replace the index set \mathbb{J}_0 with a superset having cardinality s without changing (2).

Thus, suppose $|\mathbb{J}_0| = s$ but $\mathbb{J}_0 \notin \mathcal{J}_s(\bar{x})$. Then there exist indices $j_1 \in \mathbb{J}_0$ and $j_2 \notin \mathbb{J}_0$ such that $x_{j_1}^+ < x_{j_2}^+$. In particular, $x_{j_2} = x_{j_2}^+ > 0$ and hence

$$[y]_s = \min_{j \in \mathbb{J}_0} x_j^+ \leq x_{j_1}^+ < x_{j_2}^+ = x_{j_2} \leq \max \left(\{x_j^- : j \in \mathbb{J}_0\} \cup \{x_j : j \notin \mathbb{J}_0\} \right) = [z]_1.$$

This contradicts the assumption that $[y]_s \geq [z]_1$ and we therefore conclude that $\mathbb{J}_0 \in \mathcal{J}_s(x)$. Proposition 3.2 now implies $y \in P_{\mathcal{K}_s}(x)$, and thus completes the proof. \square

We now provide our first main result: a novel characterization of the Mordukhovich normal cone to the set of non-negative sparse vectors. Given $y \in \mathbb{R}^m$ and an index set $\mathbb{J} \subseteq \{1, \dots, m\}$ the notation $y|_{\mathbb{J}} = 0$ means $y_j = 0$ for all $j \in \mathbb{J}$.

Theorem 3.4 (Mordukhovich normal cone to \mathcal{K}_s). *The Mordukhovich normal cone to the set \mathcal{K}_s at a point $\bar{x} \in \mathcal{K}_s$ is given by*

$$N_{\mathcal{K}_s}(\bar{x}) = \{y \in \mathbb{R}^m : \bar{x} \odot y = 0, y \leq 0\} \cup \{y \in \mathbb{R}^m : \bar{x} \odot y = 0, \|y\|_0 \leq m - s\}.$$

Proof. (\subseteq) Suppose $y \in N_{\mathcal{K}_s}(\bar{x})$. Then there exists sequences $(x_k), (y_k), (z_k) \subseteq \mathbb{R}^m$ such that

$$x_k \rightarrow \bar{x}, \quad y_k = \alpha_k z_k \rightarrow y, \quad z_k = x_k - p_k,$$

where $\alpha_k \in \mathbb{R}_+$ and $p_k \in P_{\mathcal{K}_s}(x_k)$. By Lemma 3.3, $p_k \odot z_k = 0$ and thus

$$x_k \odot y_k = (z_k + p_k) \odot (\alpha_k z_k) = \alpha_k (z_k \odot z_k) = z_k \odot y_k.$$

The definition of $P_{\mathcal{K}_s}$ implies that $\|z_k\| = d(x_k, \mathcal{K}_s)$. Hence, by noting that $x_k \rightarrow \bar{x} \in \mathcal{K}_s$ and that the function $d(\cdot, \mathcal{K}_s)$ is continuous, we deduce that $z_k \rightarrow 0$. Altogether

$$\begin{aligned} \bar{x} \odot y &= \left(\lim_{k \rightarrow \infty} x_k \right) \odot \left(\lim_{k \rightarrow \infty} y_k \right) = \lim_{k \rightarrow \infty} (x_k \odot y_k) \\ &= \lim_{k \rightarrow \infty} (z_k \odot y_k) \\ &= \left(\lim_{k \rightarrow \infty} z_k \right) \odot \left(\lim_{k \rightarrow \infty} y_k \right) = 0 \odot y = 0. \end{aligned}$$

By Proposition 3.2(d), for each $k \in \mathbb{N}$, there is an index set $\mathbb{J}_k \in \mathcal{J}_s(x_k)$ such that p_k is of the form

$$(p_k)_j = \begin{cases} (x_k)_j^+, & j \in \mathbb{J}_k, \\ 0, & j \notin \mathbb{J}_k. \end{cases}$$

The collection $\{\mathbb{J}_k : k \in \mathbb{N}\}$ is finite, and thus, by the Pigeonhole Principle, there exists a subsequence (k_l) and an index set $\mathbb{J}_0 \in \mathcal{J}_s(x_{k_l})$ such that

$$(p_{k_l})_j = \begin{cases} (x_{k_l})_j^+, & j \in \mathbb{J}_0, \\ 0, & j \notin \mathbb{J}_0; \end{cases} \quad (y_{k_l})_j = \begin{cases} (x_{k_l})_j^-, & j \in \mathbb{J}_0, \\ (x_{k_l})_j, & j \notin \mathbb{J}_0. \end{cases} \quad (3)$$

Since $\mathbb{J}_0 \in \mathcal{J}_s(x_{k_l})$, for all $l \in \mathbb{N}$ we have

$$\min_{j \in \mathbb{J}_0} (x_{k_l})_j^+ \geq \max_{j \notin \mathbb{J}_0} (x_{k_l})_j^+,$$

and hence

$$\min_{j \in \mathbb{J}_0} \bar{x}_j^+ = \lim_{l \rightarrow \infty} \left(\min_{j \in \mathbb{J}_0} (x_{k_l})_j^+ \right) \geq \lim_{l \rightarrow \infty} \left(\max_{j \notin \mathbb{J}_0} (x_{k_l})_j^+ \right) = \max_{j \notin \mathbb{J}_0} \bar{x}_j^+.$$

We therefore conclude that $\mathbb{J}_0 \in \mathcal{J}_s(\bar{x})$. If $y \not\leq 0$ then there is an index $j_0 \in \{1, 2, \dots, m\}$ such that $y_{j_0} > 0$. Since $y_k \rightarrow y$, we assume l to be sufficiently large so that $(y_{k_l})_{j_0} > 0$. From the representation of y_{k_l} in (3) we deduce that $j_0 \notin \mathbb{J}_0$ and $(x_{k_l})_{j_0} = (y_{k_l})_{j_0} > 0$. By the definition of \mathbb{J}_0 , it follows that

$$(x_{k_l})_j \geq (x_{k_l})_{j_0} > 0 \implies (x_{k_l})_j^- = 0,$$

for all $j \in \mathbb{J}_0$. By (3) we deduce $y_{k_l}|_{\mathbb{J}_0} = 0$. Since $y_{k_l} \rightarrow y$ it follows that $y|_{\mathbb{J}_0} = 0$, and therefore $\|y\|_0 \leq m - s$.

(\supseteq) Suppose $\bar{x} \odot y = 0$, or equivalently, for each $j \in \{1, 2, \dots, m\}$, we have that

$$\bar{x}_j \text{ and } y_j \text{ cannot be simultaneously non-zero.} \quad (4)$$

For the $y \leq 0$ case, define sequences (x_k) and (y_k) by

$$x_k := \bar{x} + \frac{1}{k} y \odot \bar{x}, \quad y_k \in k(x_k - P_{\mathcal{K}_s}(x_k)).$$

Then, by noting (4), for any $k \in \mathbb{N}$, we have $P_{\mathcal{K}_s}(x_k) = \{\bar{x}\}$ thus

$$y_k = k \left(\left(\bar{x} + \frac{y}{k} \right) - \bar{x} \right) = y \in N_{\mathcal{K}_s}(\bar{x}).$$

For the other case, suppose $\|y\|_0 \leq m - s$. Then, using (4), we see that there exists an index set $\mathbb{J}_0 \in \mathcal{J}_s(\bar{x})$ such that $y|_{\mathbb{J}_0} = 0$. Let $w \in \mathbb{R}^m$ be the vector whose entries are 1 on \mathbb{J}_0 , and 0 otherwise. Define sequences (x_k) and (y_k) by

$$x_k := \bar{x} + \frac{1}{k} y + \frac{1}{\sqrt{k}} w \rightarrow \bar{x}, \quad y_k \in k(x_k - P_{\mathcal{K}_s}(x_k)).$$

Since $1/k \rightarrow 0$ at a faster rate than $1/\sqrt{k} \rightarrow 0$, there exists a sufficiently large K such that for $k > K$,

$$\min_{j \in \mathbb{J}_0} \left\{ \bar{x}_j + \frac{1}{\sqrt{k}} w_j \right\} > \frac{1}{k} \max_{j \notin \mathbb{J}_0} \{y_j\}.$$

Hence for $k > K$ we have $P_{\mathcal{K}_s}(x_k) = \{\bar{x} + w/\sqrt{k}\}$, and thus that

$$y_k = k \left(\left(\bar{x} + \frac{1}{k} y + \frac{1}{\sqrt{k}} w \right) - \left(\bar{x} + \frac{1}{\sqrt{k}} w \right) \right) = y \in N_{\mathcal{K}_s}(\bar{x}).$$

This completes the proof. \square

For convex sets the convex and Mordukhovich normal cones coincide [27, Ch. 1]. As an immediate consequence of Theorem 3.4 we recover the following well-known result.

Corollary 3.5 (Normal cone to \mathbb{R}_+^m). *The normal cone to the convex set \mathbb{R}_+^m at the point $\bar{x} \in \mathbb{R}_+^m$ is given by*

$$N_{\mathbb{R}_+^m}(\bar{x}) = N_{\mathbb{R}_+^m}^{\text{conv}}(\bar{x}) = \{y \in \mathbb{R}^m : \bar{x} \odot y = 0, y \leq 0\}.$$

Proof. Note that $\mathbb{R}_+^m = \mathcal{K}_m$ and $\{y \in \mathbb{R}^m : \bar{x} \odot y = 0, \|y\|_0 \leq 0\} = \{0\}$. The result now follows by applying Theorem 3.4 with $s = m$. \square

The following remark sheds light on the two sets in the normal cone formula of Theorem 3.4.

Remark 3.6 ($N_{\mathcal{K}_s}$ is the union of two normal cones). The Mordukhovich normal to sparsity set \mathcal{A}_s is given by (see [5, Th. 3.9])

$$N_{\mathcal{A}_s}(\bar{x}) = \{y \in \mathbb{R}^m : \bar{x} \odot y = 0, \|y\|_0 \leq m - s\}. \quad (5)$$

Combining with Corollary 3.5, Theorem 3.4 can be expressed

$$N_{\mathcal{K}_s}(\bar{x}) = N_{\mathbb{R}_+^m}^{\text{conv}}(\bar{x}) \cup N_{\mathcal{A}_s}(\bar{x}).$$

That is, $N_{\mathcal{K}_s}$ is the union of the convex normal cone to the non-negativity set \mathbb{R}_+^m , and the Mordukhovich normal cone to the sparsity set \mathcal{A}_s . \diamond

Remark 3.7. Note that formulae for $N_{\mathbb{R}_+^m}$ and $N_{\mathcal{A}_s}$ are known, and that $\mathcal{K}_s = \mathcal{A}_s \cap \mathbb{R}_+^m$. Nevertheless, it is not possible to obtain the normal cone $N_{\mathcal{K}_s}$ using the standard *intersection rule* [27, §3.1.1] applied to $N_{\mathcal{A}_s}$ and $N_{\mathbb{R}_+^m}$ since the basic qualification condition $N_{\mathcal{A}_s}(\bar{x}) \cap (-N_{\mathbb{R}_+^m})(\bar{x}) = \{0\}$ is not satisfied. \diamond

Around points of maximal sparsity the set \mathcal{K}_s is locally indistinguishable from the sparsity set \mathcal{A}_s . In this case, the formula for the normal cone simplifies accordingly.

Corollary 3.8 (Points of maximal sparsity). *The Mordukhovich normal cone to the set \mathcal{K}_s at a point $\bar{x} \in \mathcal{K}_s$ having $\|\bar{x}\|_0 = s$ is given by*

$$N_{\mathcal{K}_s}(\bar{x}) = N_{\mathcal{A}_s}(\bar{x}) = \{y \in \mathbb{R}^m : \bar{x} \odot y = 0\}.$$

Proof. Since $\|x\|_0 = s$ we have $\{y \in \mathbb{R}^m : \bar{x} \odot y = 0, \|y\|_0 \leq m - s\} = \{y : \bar{x} \odot y = 0\}$. Observe,

$$\{y \in \mathbb{R}^m : \bar{x} \odot y = 0, y \leq 0\} \subseteq \{y \in \mathbb{R}^m : \bar{x} \odot y = 0\}.$$

The result now follows from Theorem 3.4 and (5). \square

To conclude our study of the vector setting, we give a characterization of the proximal normal cone to \mathcal{K}_s , in terms of already introduced objects.

Theorem 3.9 (Proximal normal cone to \mathcal{K}_s). *The proximal normal cone to \mathcal{K}_s at the point $\bar{x} \in \mathcal{K}_s$ is given by*

$$N_{\mathcal{K}_s}^{\text{prox}}(\bar{x}) = \begin{cases} N_{\mathbb{R}_+^m}^{\text{conv}}(\bar{x}) = N_{\mathbb{R}_+^m}(\bar{x}), & \|\bar{x}\|_0 < s, \\ N_{\mathcal{A}_s}^{\text{prox}}(\bar{x}) = N_{\mathcal{A}_s}(\bar{x}) = N_{\mathcal{K}_s}(\bar{x}), & \|\bar{x}\|_0 = s. \end{cases}$$

Proof. On one hand, if $\|\bar{x}\|_0 < s$ then, by Proposition 3.2, $P_{\mathcal{K}_s}^{-1}(\bar{x}) = P_{\mathbb{R}_+^m}^{-1}(\bar{x})$ implying that the corresponding proximal normal cones coincide. The claimed formula now follows by Fact 2.1. On the other hand, if $\|\bar{x}\|_0 = s$ the result follows from Corollary 3.8 and [5, Prop. 3.8]. \square

Remark 3.10. Theorem 3.9 shows that the proximal normal cone to a non-negative sparsity set does not capture all important features of the set at points which do not have maximal sparsity. In particular, the proximal normal cone coincides with the normal cone to the convex non-negative cone, whereas the corresponding Mordukhovich cone need not. Similar behavior is observed for the proximal normal cone to the sparsity set \mathcal{A}_s , which is equal to $\{0\}$ at points not having maximal sparsity [5, Prop. 3.8]. \diamond

3.2 Low-Rank Positive Semi-Definite Matrices

Using the correspondence between symmetric and spectral functions, we now *lift* our vector results to the larger space of symmetric matrices.

Proposition 3.11 (Projection onto \mathcal{S}_s). *Let $X \in \mathbb{S}^m$ and define*

$$\lambda_s^+(X) := (\lambda_1^+(X), \dots, \lambda_s^+(X), 0, \dots, 0).$$

The projection of X onto \mathcal{S}_s is given by

$$P_{\mathcal{S}_s}(X) = \{Y \in \mathbb{S}^m : Y = U^\top (\text{diag } \lambda_s^+(X)) U, U \in \mathbb{O}^m(X)\}.$$

Proof. Follows from Fact 2.3 and Proposition 3.2. \square

Our next main result is a novel characterization of the Mordukhovich normal cone to the set \mathcal{S}_s .

Theorem 3.12 (Mordukhovich normal cone to \mathcal{S}_s). *The Mordukhovich normal cone to the set \mathcal{S}_s at the point $\bar{X} \in \mathcal{S}_s$ is given by*

$$N_{\mathcal{S}_s}(\bar{X}) = \{Y \in \mathbb{S}^m : \bar{X}Y = 0, Y \preceq 0\} \cup \{Y \in \mathbb{S}^m : \bar{X}Y = 0, \text{rank}(Y) \leq m - s\}.$$

Proof. First observe that, by combining Theorem 3.4 and Fact 2.4, we obtain

$$N_{\mathcal{S}_s}(\overline{X}) = \{U^\top \text{diag}(y)U : \lambda(\overline{X}) \odot y = 0, y \leq 0, U \in \mathbb{O}^m(X)\} \\ \cup \{U^\top \text{diag}(y)U : \lambda(\overline{X}) \odot y = 0, \|y\|_0 \leq m - s, U \in \mathbb{O}^m(X)\}. \quad (6)$$

To complete the proof, we show that (6) is equal to the claimed formula. Since other inclusion is easily deduced, we only prove that any matrix Y satisfying the proposed formula is contained in (6). To this end, consider $Y \in \mathbb{S}^m$ with $\overline{X}Y = 0$. Then \overline{X} and Y commute, and hence have a simultaneously spectral decomposition which we assume, without loss of generality, is an ordered spectral decomposition for \overline{X} . That is, there exists $U \in \mathbb{O}^m(\overline{X})$ such that

$$Y = U^\top \text{diag}(y)U \text{ for some } y \in \mathbb{R}^m.$$

Furthermore, since $U \in \mathbb{O}^m$, we have that

$$\begin{aligned} \text{diag}(\lambda(\overline{X}) \odot y) &= \text{diag}(\lambda(\overline{X})) \text{diag}(y) \\ &= U (U^\top \text{diag}(\lambda(\overline{X}))U) (U^\top \text{diag}(y)U) U^\top \\ &= U \overline{X} Y U^\top = 0, \end{aligned}$$

which implies that $\lambda(X) \odot y = 0$. To complete the proof, we note that $Y \preceq 0$ implies $y \leq 0$ and that $\text{rank } Y \leq m - s$ implies $\|y\|_0 \leq m - s$. \square

As before, we deduce consequences of Theorem 3.12. The first is the normal cone to the set of positive semi-definite matrices. This can be found, for example, in [20].

Corollary 3.13 (Normal cone to \mathbb{S}_+^m). *The normal cone to the set \mathbb{S}_+^m at a point $\overline{X} \in \mathbb{S}_+^m$ is given by*

$$N_{\mathbb{S}_+^m}(\overline{X}) = N_{\mathbb{S}_+^m}^{\text{conv}}(\overline{X}) = \{Y \in \mathbb{S}^+ : \overline{X}Y = 0, Y \preceq 0\}.$$

Proof. Note that $\mathbb{S}_+^m = \mathcal{S}_s$ and $\{Y \in \mathbb{S}^m : \overline{X}Y = 0, \text{rank}(Y) \leq 0\} = \{0\}$. The result now follows from Theorem 3.12 with $s = m$. \square

Denote the set of low-rank symmetric matrices by

$$\mathcal{R}_s := \{X \in \mathbb{S}^m : \text{rank}(X) \leq s\}.$$

The following proposition is a characterization of the Mordukhovich normal cone to \mathcal{R}_s .

Proposition 3.14 (Mordukhovich normal cone to \mathcal{R}_s). *The Mordukhovich normal cone to the set \mathcal{R}_s at a $\overline{X} \in \mathcal{R}_s$ having $\text{rank } \overline{X} = s$ is given by*

$$N_{\mathcal{R}_s}(\overline{X}) = \{Y \in \mathbb{S}^m : \overline{X}Y = 0\}.$$

Proof. Follows from (5) and Fact 2.4. \square

Remark 3.15. A formula for the Mordukhovich normal cone to the set of low-rank (possibly rectangular) real matrices is derived in [26]. This formula cannot be applied in our context even, when specialized to the square case, for the same reasons discussed in Remark 3.7 \diamond

As was the case in the vector setting, here we are also able to express the normal cone to \mathcal{S}_s in terms of the normal cones of its two ‘building blocks’.

Remark 3.16 ($N_{\mathcal{S}_s}$ is the union of two normal cones). In light of Theorem 3.12, Corollary 3.13 and Proposition 3.14, the Mordukhovich normal cone to \mathcal{S}_s can be expressed as

$$N_{\mathcal{S}_s}(\overline{X}) = N_{\mathbb{S}_+^m}^{\text{conv}}(\overline{X}) \cup N_{\mathcal{R}_s}(\overline{X}).$$

That is, $N_{\mathcal{S}_s}$ is the union of the convex normal cone to the positive semi-definite matrices, and the Mordukhovich normal cone to the low-rank set \mathcal{R}_s . \diamond

A characterization of the proximal normal cone to \mathcal{S}_s can also be given. As was the case in the vector setting, it also does not adequately describe the geometry of \mathcal{S}_s .

Theorem 3.17 (Proximal normal cone to \mathcal{S}_s). *The proximal normal cone to \mathcal{S}_s at the point $\overline{X} \in \mathcal{K}_s$ is given by*

$$N_{\mathcal{S}_s}^{\text{prox}}(\overline{X}) = \begin{cases} N_{\mathbb{S}_+^m}^{\text{conv}}(\overline{X}) = N_{\mathbb{S}_+^m}(\overline{X}), & \text{rank}(\overline{X}) < s, \\ N_{\mathcal{R}_s}^{\text{prox}}(\overline{X}) = N_{\mathcal{R}_s}(\overline{X}), & \text{rank}(\overline{X}) = s. \end{cases}$$

Proof. Follows from Fact 2.4 and Theorem 3.9. \square

4 Regularity Properties

We now investigate regularity properties of non-negative sparsity sets, and collections of sets containing non-negative sparsity sets. We first recall some definitions.

A family $\{\Omega_1, \Omega_2, \dots, \Omega_n\}$ of closed non-empty subsets of \mathbb{E} is *strongly regular* at a point $\overline{x} \in \bigcap_{k=1}^n \Omega_k$ if $y_1 = y_2 = \dots = y_m = 0$ is the only solution to the system

$$\sum_{k=1}^m y_k = 0, \quad y_k \in N_{\Omega_k}(\overline{x}).$$

Specialized to a collection of two sets, strong regularity can be expressed as the transversality condition

$$N_{\Omega_1}(\overline{x}) \cap (-N_{\Omega_2}(\overline{x})) = \{0\}. \quad (7)$$

The normal cone formulae in Theorems 3.4 and 3.12 provide conditions for strongly regularity of intersections involving non-negative sparsity sets. These condition, which we derive below, are analogous to those given by [26, Prop. 3.8] for rank constraints not requiring non-negativity.

Proposition 4.1 (Strong regularity for intersections with \mathcal{K}_s). *Let $\Omega \subseteq \mathbb{R}^m$ be closed and $\overline{x} \in \Omega \cap \mathcal{K}_s$. Then $\{\Omega, \mathcal{K}_s\}$ is strongly regular at \overline{x} if and only if for all non-zero $y \in N_{\Omega}(\overline{x})$ either (a) $\overline{x} \odot y \neq 0$, or (b) $y \not\geq 0$ and $\|y\|_0 > m - s$.*

Proof. By characterization of $N_{\mathcal{K}_s}$ in Theorem 3.4, we deduce that strong regularity of the sets $\{\Omega, \mathcal{K}_s\}$ is equivalent to

$$\begin{aligned} \{0\} &= N_{\Omega}(\overline{x}) \cap \{y \in \mathbb{R}^m : \overline{x} \odot y = 0, y \geq 0\}, \text{ and} \\ \{0\} &= N_{\Omega}(\overline{x}) \cap \{y \in \mathbb{R}^m : \overline{x} \odot y = 0, \|y\|_0 \leq (m - s)\}. \end{aligned}$$

The result now follows. \square

The following Proposition is the symmetric matrix analogue of Proposition 4.1.

Proposition 4.2 (Strong regularity for intersections with \mathcal{S}_s). *Let $\Omega \subseteq \mathbb{S}^m$ be closed and $\bar{X} \in \Omega \cap \mathcal{S}_s$. Then $\{\Omega, \mathcal{S}_s\}$ is strongly regular at \bar{X} if and only if for all non-zero $Y \in N_\Omega(\bar{X})$ either (a) $\bar{X}Y \neq 0$, or (b) $Y \not\preceq 0$ and $\text{rank}(Y) > m - s$.*

Proof. Argue similarly to Proposition 4.1, using the characterization of $N_{\mathcal{S}_s}$ in Theorem 3.12. \square

In Section 5.1, we give an application of Proposition 4.1 to solving sparse linear systems, and of Proposition 4.2 to low-rank semi-definite programming feasibility.

Remark 4.3 (Affine-hull regularity). *Affine-hull regularity* is a regularity notion for collections of sets which is weaker than strong regular which has been utilized to obtain convergence results in the absence of strong regularity [6, 29]. However, as we now explain, for collections of constraint sets containing a sparsity set, the two notions coincide.

Focusing on the case of two sets, affine-hull regularity can be viewed as a modification of the definition of strong regularity in which the Mordukovich normal cones in (7) are replaced with the so-called *restricted Mordukovich normal cones*, developed in [6, 5], with the restriction performed with respect to the affine-hull of the union of the two constraint sets. If $s \geq 1$ and the affine-hull of any of sparsity sets \mathcal{K}_s , \mathcal{A}_s , \mathcal{S}_s or \mathcal{R}_s is equal to the entire space. The corresponding affine restricted Mordukovich normal and standard Mordukovich normal cones coincide and, it therefore follows that, for a collection of constraint sets containing such a sparsity set, the notions of strong and affine-hull regularity coincide. \diamond

A closed set $\Omega \subseteq \mathbb{E}$ is *prox-regular* at $\bar{x} \in \Omega$ if for all $\bar{y} \in N_\Omega(\bar{x})$ there exists an $\epsilon > 0$ and $\rho > 0$ such that whenever $\|x - \bar{x}\| < \epsilon$ and $y \in N_\Omega(x)$ with $\|y - \bar{y}\| < \epsilon$ then x is the unique nearest point of $\{z \in \Omega : \|z - \bar{x}\| < \epsilon\}$ to $x + \rho y$. A useful characterization which we utilize in the following proofs is that prox-regularity at $\bar{x} \in \Omega$ is equivalent to single-valuedness of P_Ω around \bar{x} [30, Th. 1.3].

Proposition 4.4 (Prox-regularity of \mathcal{K}_s). *Let $\bar{x} \in \mathcal{K}_s$ with $s \in \{1, \dots, (m - 1)\}$ and $m \geq 2$. Then \mathcal{K}_s is prox-regular at \bar{x} if and only if $\|\bar{x}\|_0 = s$.*

Proof. On one hand, suppose $\|\bar{x}\|_0 < s$. To show that \mathcal{K}_s is not prox-regular at \bar{x} , it suffices to produce a sequence $x_k \rightarrow \bar{x}$ such that $P_{\mathcal{K}_s}(x_k)$ is not singleton for any $k \in \mathbb{N}$. To this end, since $\|\bar{x}\|_0 \leq (s - 1)$, we have

$$2 \leq (s - \|\bar{x}\|_0) + 1 = (s + 1) - \|\bar{x}\|_0 \leq m - \|\bar{x}\|_0.$$

Since x has $(m - \|x\|_0)$ entries which are zero, therefore there exists an index set $\mathbb{I}_0 \subseteq \{1, \dots, m\}$ with $|\mathbb{I}_0| = (s - \|\bar{x}\|_0) + 1 \geq 2$ such that $\bar{x}_j = 0$ for all $j \in \mathbb{I}_0$. Define the vector $v \in \mathbb{R}^m$ and the sequence (x_k) by

$$v := \sum_{j \in \mathbb{I}_0} e_j, \quad x_k := \bar{x} + \frac{1}{k}v.$$

Observe $\|x_k\|_0 = \|\bar{x}\|_0 + \|v\|_0 = s + 1$. Thus, for sufficiently large k , Proposition 3.2 yields

$$P_{\mathcal{K}_s}(x_k) = \left\{ \bar{x} + \frac{1}{k}(v - e_j) : j \in \mathbb{I}_0 \right\} \text{ where } |P_{\mathcal{K}_s}(x_k)| = |\mathbb{I}_0| \geq 2.$$

By taking k sufficiently large, the distance between x_k and \bar{x} can be made arbitrary small, thus proving that \mathcal{K}_s is not prox-regular at \bar{x} .

On the other hand, suppose $\|\bar{x}\|_0 = s$, and let $\delta := \frac{1}{2} \min\{\bar{x}_j : \bar{x}_j > 0\} > 0$. Since \bar{x} has maximal sparsity, the set $\mathcal{J}_s(\bar{x})$ is a singleton, say $\mathcal{J}_s(\bar{x}) = \{\mathbb{J}_0\}$. For any $x \in \mathbb{B}_\delta(\bar{x})$, $j \in \mathbb{J}_0$ and $i \notin \mathbb{J}_0$ we have

$$\bar{x}_j - x_j \leq |\bar{x}_j - x_j| \leq \|\bar{x} - x\| < \delta \leq \frac{1}{2}\bar{x}_j, \quad \text{and} \quad x_i = |x_i - \bar{x}_i| \leq \|\bar{x} - x\| < \delta \leq \frac{1}{2}\bar{x}_j.$$

Altogether, $0 < \bar{x}_j/2 \leq x_j$, $x_i < \bar{x}_j/2$, and therefore $\mathcal{J}_s(x) = \{\mathbb{J}_0\}$. By Proposition 3.2 $P_{\mathcal{K}_s}(x)$ is single-valued, hence \mathcal{K}_s is prox-regular at \bar{x} . \square

Proposition 4.5 (Prox-regularity of \mathcal{S}_s). *Let $\bar{X} \in \mathcal{S}_s$ with $s \in \{1, \dots, (m-1)\}$ and $m \geq 2$. Then \mathcal{S}_s is prox-regular at \bar{X} if and only if $\text{rank}(\bar{X}) = s$.*

Proof. By Proposition 4.4, the set \mathcal{K}_s is prox-regular at $\lambda(\bar{X})$ if and only if $\|\lambda(\bar{X})\|_0 = s$. The result now follows from [11, Th. 9]. \square

Remark 4.6. In both of the above propositions, we excluded the possibility that $s = m$. However, in this case, the sets are convex hence everywhere prox-regular. \diamond

5 Implications and Applications

In this section we give instances of problems to which the derived normal cone formulae and regularity results apply. In particular, we consider instances of the (2-set) *feasibility problem*. That is, given sets $C_1, C_2 \subseteq \mathbb{E}$, we aim to

$$\text{find } x \in C_1 \cap C_2 \neq \emptyset. \tag{8}$$

The study of these regularity properties are of interest, in part, because they are precisely the ingredient required by the current state-of-art theory for non-convex *projection algorithms* to guarantee local convergence [5, 6, 19, 29]. Examples of such algorithms include *cyclic projection* and *Douglas–Rachford methods*. These algorithms are intimately related to a number fundamental algorithms arising the optics literature including the *error reduction*, the *difference-map* and *hybrid input-output* algorithms [4, 16, 15].

In each of the following examples, we first state the problem and then give an appropriate feasibility formulation or relaxation. The conditions required for the regularity properties studied to hold are then considered. Particular attention is given to our final example which arise in low-rank distance matrix reconstruction problems. In this direction, we given a characterization of the sufficient conditions required by [29] to guarantee local linear convergence of the *Douglas–Rachford method*. This result complements the empirical studies [9, 1] in which local linear convergence was observed. Moreover, to our knowledge, there are no other results in the literature which provide sufficient conditions for convergence of the Douglas–Rachford algorithm applied to this problem.

5.1 Sparse and Low-Rank Linear Systems

Consider the problem of finding the sparsest non-negative solution to a linear system. That is,

$$\min_{x \in \mathbb{R}^n} \{\|x\|_0 : Ax = b, x \geq 0\}, \quad (9)$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

When a desired sparsity bound, s , can be given *a priori*, the problem can be reformulated as

$$\text{find } x \in \mathbb{R}^m \text{ such that } Ax = b, x \geq 0, \|x\|_0 \leq s. \quad (10)$$

As we have already mentioned, problems of precisely this kind arise when a signal x , which is known *a priori* to be non-negative, purely real and sparse, is to be recovered from the under-determined linear system specified by the measurement matrix A and observation vector b . Another example is given by *mixed-integer linear programs*.

Example 5.1 (Mixed-binary linear program). Let $M > 0$ and consider a mixed-binary linear program having feasible region F given by

$$F := \left\{ (x, y) \in \mathbb{R}^m \times \{0, 1\}^m : Ax = b, 0 \leq x \leq My, \sum_{j=1}^m y_j \leq s \right\}.$$

Finding a pair contained in F can be formulated in terms of (10). Precisely,

$$(x, y) \in F \iff \begin{cases} x \in \mathbb{R}^m \text{ satisfies (10) and } \|x\|_\infty \leq M, \\ y \in \{0, 1\}^m \text{ with } y_j = 1 \text{ whenever } x_j \neq 0. \end{cases}$$

Regarding this equivalence, note that in order to find a pair $(x, y) \in F$, it suffices to find an $x \in \mathbb{R}^m$ satisfying (10) with $\|x\|_\infty \leq M$. Using such a point, a feasible vector $y \in \{0, 1\}^m$ can easily be computed since $\|x\|_0 \leq s$.

Finding feasible points of mixed-integer linear programs is important as they are frequently solved using *branch-and-bound* techniques which require feasible solutions during *pruning steps*. \diamond

Returning our attention to the general problem, we observe (10) to be equivalent to the feasibility problem (8) with constraint sets

$$\begin{aligned} C_1 &:= \{x \in \mathbb{R}^m : Ax = b\}, \\ C_2 &:= \mathcal{K}_s = \{x \in \mathbb{R}^n : x \geq 0, \|x\|_0 \leq s\}. \end{aligned} \quad (11)$$

We may now give a characterisation of strong regularity of the collection $\{C_1, C_2\}$ using the results of previous sections.

Proposition 5.2. *Suppose $\bar{x} \in C_1 \cap C_2$ where the constraint sets C_1 and C_2 are as defined in (11). Then $\{C_1, C_2\}$ is strongly regular at \bar{x} if and only if, for all $y \in \text{range } A^\top \setminus \{0\}$, either $\bar{x} \odot y = 0$, or (b) $y \not\geq 0$ and $\|y\|_0 > m - s$.*

Proof. Suppose $\bar{x} \in C_1 \cap C_2$. The normal cone to C_1 at \bar{x} is given by (see [7, Sec. 2.1, Exer. 4])

$$N_{C_1}(\bar{x}) = \text{range } A^\top.$$

The result now follows from Proposition 4.1 with $\Omega := C_1$. \square

We now turn our attention to the symmetric matrix analogue. Consider the problem of finding a minimum rank positive semi-definite matrix solution to a linear system. That is,

$$\min_{X \in \mathbb{S}^n} \{\text{rank } X : (\langle A_j, X \rangle)_{j=1}^p = b, X \succeq 0\}. \quad (12)$$

As before, when a desired bound on the rank, s , can be given *a priori*, the problem can be reformulated as

$$\text{find } X \in \mathbb{S}^m \text{ such that } (\langle A_j, X \rangle)_{j=1}^p = b, X \succeq 0, \text{rank}(X) \leq s, \quad (13)$$

where $A_j \in \mathbb{S}^m$ for each $j \in \{1, \dots, p\}$ and $b \in \mathbb{R}^p$.

Problems of this kind arise in rank-constrained semi-definite programming [21, Exer. 2.4], an NP-hard problem.

We observe (13) to be equivalent to the two set feasibility problem (8) with constraint sets

$$\begin{aligned} C_1 &:= \{X \in \mathbb{S}^m : (\langle A_j, X \rangle)_{j=1}^p = b\}, \\ C_2 &:= \mathcal{S}_s = \{X \in \mathbb{S}^m : X \succeq 0, \text{rank } X \leq s\}. \end{aligned} \quad (14)$$

We may now give a characterization of strong regularity of the collection $\{C_1, C_2\}$.

Proposition 5.3. *Suppose $\bar{X} \in C_1 \cap C_2$ where the constraint sets C_1 and C_2 are as defined in (14). Then $\{C_1, C_2\}$ is strongly regular at \bar{X} if and only if, for all $Y \in \text{span}\{A_1, A_2, \dots, A_p\} \setminus \{0\}$, either (a) $\bar{X}Y = 0$, or (b) $Y \not\preceq 0$ and $\text{rank}(Y) > m - s$.*

Proof. Suppose $\bar{X} \in C_1 \cap C_2$. Then the normal cone to C_1 at \bar{X} is given by (see [7, Sec 2.1, Exer 4] & [2, Sec. 5])

$$N_{C_1}(\bar{X}) = \text{range} \left((\langle A_j, \cdot \rangle)_{j=1}^p \right)^* = \text{span}\{A_1, A_2, \dots, A_p\}.$$

The result now follows from 4.2 with $\Omega := C_1$. \square

Remark 5.4 (Sparsity upper bounds). In order to apply Formulations (10) and (13), it is necessary that the sparsity/rank parameter s is known *a priori*. While this is not always possible, we emphasise that it is usually not the case that the *optimal* sparsity parameter (in the sense of (9) or (12)) is needed. When projection or reflection methods are applied to a formulation only prescribing an upper bound for the parameter, there is nothing to prevent sparser/lower-rank solutions from being returned. For further details, see [19].

5.2 Low-Rank Euclidean Distance Matrix Completion

We now use the derived normal cone formulae to provide theoretical justification for an application of the *Douglas–Rachford method*. Given constraint sets C_1 and C_2 with non-empty intersection, and initial point x_0 , this method generates a sequence (x_n) by selecting

$$x_{n+1} \in Tx_n := \left(\frac{I + R_{C_2}R_{C_1}}{2} \right) x_n.$$

Here $R_\Omega := 2P_\Omega - I$ denotes the *reflection mapping w.r.t.* Ω . If the sequence (x_n) converges to a fixed point $\bar{x} \in \text{Fix } T := \{x : x \in Tx\}$, then there is an element $\bar{p} \in P_{C_1}(\bar{x})$ such

that $\bar{p} \in C_1 \cap C_2$. That is, the point \bar{p} solves the *feasibility problem* (rather than the fixed point \bar{x} itself). For closed convex sets, this method is well-understood and global convergence guaranteed [3].

Recall a *Euclidean distance matrix* (EDM) is a matrix $D = (D_{ij}) \in \mathbb{S}^{m+1}$ such that there exists points $p_1, p_2, \dots, p_m, p_{m+1} \in \mathbb{R}^q$ such that

$$D_{ij} = \|p_i - p_j\|^2 \text{ for all } (i, j) \in \{1, 2, \dots, (m+1)\} \times \{1, 2, \dots, (m+1)\}. \quad (15)$$

Clearly any EDM is non-negative and hollow (*i.e.*, contains only zeros along its main diagonal). When (15) holds for a set points in \mathbb{R}^q , we say D is *embeddable* in \mathbb{R}^q . If D is embeddable in \mathbb{R}^q but not \mathbb{R}^{q-1} then we say D is *irreducibly embeddable* in \mathbb{R}^q .

In [1, 9] the authors consider the problem of reconstructing a *low-rank Euclidean distance matrix*, X , knowing only a subset of its entries. Let $D = (D_{ij})$ be the partial EDM with the position of the known entries specified by the index set \mathcal{I} . Assuming the EDM is *embeddable* in \mathbb{R}^s , this problem can be formulated as the feasibility problem (8) with constraint sets

$$\begin{aligned} C_1 &= \{X \in \mathbb{S}^{m+1} : X \geq 0, X_{ij} = D_{ij} \text{ for all } (i, j) \in \mathcal{I}\}, \\ C_2 &= \left\{X \in \mathbb{S}^{m+1} : Q(-X)Q = \begin{bmatrix} \hat{X} & d \\ d^\top & \delta \end{bmatrix}, \hat{X} \in \mathcal{S}_s, d \in \mathbb{R}^m, \delta \in \mathbb{R}\right\}. \end{aligned} \quad (16)$$

Here the matrix $Q \in \mathbb{O}^{m+1}$ is the *Householder matrix* given by

$$Q = I - \frac{2vv^\top}{v^\top v}, \text{ where } v = (1, 1, \dots, 1, 1 + \sqrt{m+1})^\top \in \mathbb{R}^{m+1}.$$

We denote by $G : \mathbb{S}^{m+1} \rightarrow \mathbb{S}^{m+1}$ the linear isometry $X \mapsto Q(-X)Q$. Further observe that $G^2 = I$. This feasibility problems is a consequence of the following EDM characterization. For further details, the reader is referred to [1, 9, 18, 31].

Theorem 5.5 (EDM characterization [18, Th. 3.3]). *Let $X \in \mathbb{S}^{m+1}$ be a non-negative, hollow matrix and denote by $\hat{X} \in \mathbb{S}^m$ the upper-left block of the matrix $Q(-X)Q$ (see (16)). Then X is a Euclidean distance matrix if and only if $\hat{X} \in \mathbb{S}_+^m$. Moreover, when X is an EDM, it is irreducibly embeddable in \mathbb{R}^s where $s = \text{rank}(\hat{X}) \leq m$.*

Remark 5.6 (Properties of the constraint sets). Observe that the constraint set C_1 is closed and convex. In fact, it is the intersection of a closed convex cone and a closed affine subspace. The constraint set C_2 is non-convex and can alternatively be described as the pre-image of the set \mathcal{S}_s under the linear operator which maps $X \in \mathbb{S}^{m+1}$ to $\hat{X} \in \mathbb{S}^m$ as given in (16). Note that both the mathematical challenging and physically meaning cases arise when the sparsity bound satisfies $1 < s < m$. \diamond

The goal of the remainder of this section is to provide conditions for which the following local convergence result due to Phan [29] applies. To apply the result we shall need to check two properties: *super-regularity* and *strong regularity*.

Theorem 5.7 (Local convergence [29, Th. 4.3]). *Suppose Ω_1 and Ω_2 are super-regular at $\bar{x} \in \Omega_1 \cap \Omega_2$ and $\{\Omega_1, \Omega_2\}$ is strongly regular at \bar{x} . For any x_0 sufficiently close to \bar{x} , the Douglas–Rachford method converges to a point $\Omega_1 \cap \Omega_2$ with R -linear rate.*

Rather than verifying super-regularity directly, we instead consider the case in which the constraints in (16) satisfy a stronger property. Later we shall see that for the application considered, this is in fact the case. We also note that, in light of Remark 4.3, there is nothing to gain in applying [29, Th. 4.7] (the affine-hull regularity analogue of Theorem 5.7).

Proposition 5.8 (Prox-regularity of constraints). *Suppose $\bar{X} \in C_1 \cap C_2$ with $\text{rank } \hat{X} = s$ where C_1, C_2 and \hat{X} are as defined in (16). Then C_1 and C_2 are prox-regular (and hence also super-regular) at \bar{X} .*

Proof. Since $\bar{X} \in C_1 \cap C_2$, both C_1 and C_2 are nonempty. As noted in Remark 5.6, C_1 is closed and convex, and hence is everywhere prox-regular. To deduce the prox-regularity of C_2 , first observe that, by Proposition 4.5 and Remark 4.6, the set \mathcal{S}_s is prox-regular at \hat{X} , and hence so too is the set

$$G(C_2) = \left\{ \begin{bmatrix} \hat{Z} & d \\ d^\top & \delta \end{bmatrix} : \hat{Z} \in \mathcal{S}_s, d \in \mathbb{R}^m, \delta \in \mathbb{R} \right\}$$

is prox-regular at $G(\bar{X})$. Since G is an isometry and $G^2 = I$, we deduce that the set $C_2 = G(G(C_2))$ is prox-regular at the point $\bar{X} = G(G(\bar{X}))$. The fact that prox-regularity implies super-regularity can be found in [24, Prop. 4.9]. \square

We now turn our attention to strong regularity of $\{C_1, C_2\}$ at \bar{X} . The next two Propositions give the respective normal cones.

Proposition 5.9 (Normal cone to C_1). *Let $\bar{X} \in C_1$. Then*

$$N_{C_1}(\bar{X}) = \{Y + Z : \bar{X} \odot Z = 0, Z \leq 0, Y_{ij} = 0 \text{ for all } (i, j) \notin \mathcal{I}\}.$$

In particular, if \bar{X} has zeros on the main diagonal, has strictly positive entries elsewhere, and $(j, j) \in \mathcal{I}$ for $j \in \{1, 2, \dots, m\}$ then

$$N_{C_1}(\bar{X}) = \{Y : Y_{ij} = 0 \text{ for all } (i, j) \notin \mathcal{I}\}.$$

Proof. Observe that C_1 is the intersection of the polyhedral sets

$$A := \{X \in \mathbb{S}^{m+1} : X_{ij} = D_{ij} \text{ for all } (i, j) \in \mathcal{I}\}, \quad K := \{X \in \mathbb{S}^{m+1} : X \geq 0\}.$$

We note that the intersection of these sets is non-empty since C_1 is assumed nonempty. The (polyhedral) sum rule [7, Cor. 5.1.9] thus ensures that

$$N_{C_1}(\bar{X}) = N_A(\bar{X}) + N_K(\bar{X}).$$

The normal cones to A and K are given by

$$\begin{aligned} N_A(\bar{X}) &= \{Y \in \mathbb{S}^{m+1} : Y_{ij} = 0 \text{ for all } (i, j) \notin \mathcal{I}\}, \\ N_K(\bar{X}) &= \{Z \in \mathbb{S}^{m+1} : \bar{X} \odot Z = 0, Z \leq 0\}, \end{aligned}$$

and the first claim follows.

In particular, suppose \overline{X} has zeros on the main diagonal and strictly positive entries elsewhere, and that $(j, j) \in \mathcal{I}$ for $j \in \{1, 2, \dots, m\}$. Let $Z \in N_K(\overline{X})$. Since $\overline{X} \odot Z = 0$, it follows that Z can be non-zero only on its main diagonal. In this case, we therefore have the inclusion

$$N_K(\overline{X}) \subseteq \{Z \in \mathbb{S}^{m+1} : Z \leq 0, Z_{ij} = 0 \text{ for all } (i, j) \notin \mathcal{I}\} \subseteq N_A(\overline{X}).$$

As before, the result follows by an application of the (polyhedral) sum rule. \square

Proposition 5.10 (Normal cone to C_2). *Let $\overline{X} \in C_2$ and $\hat{X} \in \mathcal{S}_s$. Then*

$$N_{G(C_2)} \left(G(\overline{X}) \equiv \begin{bmatrix} \hat{X} & d \\ d^\top & \delta \end{bmatrix} \right) = \left\{ \begin{bmatrix} \hat{Y} & 0 \\ 0 & 0 \end{bmatrix} : \hat{Y} \in N_{\mathcal{S}_s}(\hat{X}) \right\}.$$

In particular, if $\text{rank } \hat{X} = s$ then

$$N_{G(C_2)} \left(G(\overline{X}) \equiv \begin{bmatrix} \hat{X} & d \\ d^\top & \delta \end{bmatrix} \right) = \left\{ \begin{bmatrix} \hat{Y} & 0 \\ 0 & 0 \end{bmatrix} : \hat{X}\hat{Y} = 0 \right\}.$$

Proof. The first formula follows from Theorem 3.12 and the fact that the normal cone of a Cartesian product is the Cartesian product of the normal cones [27, Prop. 1.2]. The second formula follows from Proposition 3.14. \square

The following proposition gives a formulation of strong regularity in terms of the normal cone formula in Proposition 5.9 & 5.10. That is, in terms of known objects.

Lemma 5.11 (Strong regularity of $\{C_1, C_2\}$). *Let $\overline{X} \in C_1 \cap C_2$. Then $\{C_1, C_2\}$ is strong regularity at \overline{X} if and only if*

$$G(N_{C_1}(\overline{X})) \cap -N_{G(C_2)}(G(\overline{X})) = \{0\}.$$

Proof. Since G is self-inverse (hence self-adjoint) and linear, applying [27, Th. 1.17] yields

$$N_{C_2}(\overline{X}) = G(N_{G(C_2)}(G(\overline{X}))) \implies G(N_{C_2}(\overline{X})) = N_{G(C_2)}(G(\overline{X})).$$

Hence

$$\begin{aligned} \{0\} = N_{C_1}(\overline{X}) \cap -N_{C_2}(\overline{X}) &\iff \{0\} = G(N_{C_1}(\overline{X})) \cap -G(N_{C_2}(\overline{X})) \\ &\iff \{0\} = G(N_{C_1}(\overline{X})) \cap -N_{G(C_2)}(G(\overline{X})). \end{aligned}$$

The proof is complete. \square

The following results shows that strong regularity of $\{C_1, C_2\}$ can be expressed in terms of a conditions which could, in principle, be checked once a solution is known.

Theorem 5.12 (Regularity and local convergence). *Suppose $\overline{X} \in C_1 \cap C_2$ is irreducibly embeddable in \mathbb{R}^s , has zeros on the main diagonal and has strictly positive entries elsewhere. Then both C_1 and C_2 are prox-regular at \overline{X} . The collection $\{C_1, C_2\}$ is strongly regular at \overline{X} if and only if there exists no non-zero pair $(Y, \hat{Y}) \in \mathbb{S}^{m+1} \times \mathbb{S}^m$ such that*

$$G(Y) = \begin{bmatrix} \hat{Y} & 0 \\ 0 & 0 \end{bmatrix}, \hat{X}\hat{Y} = 0, Y_{ij} = 0 \text{ for all } (i, j) \notin \mathcal{I}. \quad (17)$$

Consequently, whenever (17) holds, the Douglas–Rachford method converges with R -linear rate whenever the initial point is sufficiently close to \overline{X} .

Proof. Prox-regularity follows from the irreducible embeddability assumption combined with Theorem 5.5 and Proposition 5.8. Equation (17) follows immediately by combining the normal cone formulae in Propositions 5.9 & 5.10 with the equivalent definition of strong regularity in Lemma 5.11. Local convergence of the Douglas–Rachford algorithm now follows from Theorem 5.7. \square

6 Conclusion

In this paper, novel formulae for the Mordukhovich normal cones to the sets of non-negative sparse vector and low-rank positive semi-definite matrix have been provided. These normal cones are of interest, for instance, in examining the regularity properties of constraint sets in various applications. As a concrete application of our results, the precise conditions under which the current non-convex state-of-the-art convergence theory for the Douglas–Rachford algorithm holds, when applied to low-rank Euclidean distance matrix reconstruction, have been characterized. Regarding this particular application, an avenue for further investigation would be to numerically check condition (17), and thus determine if strong regularity holds in real-world datasets. It is worth noting that, in (17), the matrix \widehat{X} and index set \mathcal{I} are specific to each problem instance. In practice, these are determined from collected data and will vary from experiment to experiment. For this reasons, it would there also be useful to see if a probabilistic argument can be used to show that the strong regularity condition in Theorem 5.12 holds in some *generic* sense.

Acknowledgements. The author would like to thank Jonathan Borwein for his suggestions. The author is supported by the Deutsche Forschungsgemeinschaft Research Training Grant 2088. The work was partly performed during the author’s candidature at the University of Newcastle with the support of an Australian Postgraduate award.

References

- [1] F.J. Aragón Artacho, J.M. Borwein, M.K. Tam, Douglas–Rachford feasibility methods for matrix completion problems, ANZIAM J. 55 (4) (2014) 299–326.
- [2] H.H. Bauschke, J.M. Borwein, On the convergence of von Neumann’s alternating projection algorithms for two sets, Set-Valued Anal. 1 (1993) 185–212.
- [3] H.H. Bauschke, P.L. Combettes, Convex analysis and monotone operator theory in Hilbert spaces, Springer, New York, 2001.
- [4] H.H. Bauschke, P.L. Combettes, D.R. Luke, Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization, JOSA A, 19 (2002) 1334–1345.
- [5] H.H. Bauschke, D.R. Luke, H.M. Phan, X. Wang, Restricted normal cones and sparsity optimization with affine constraints, Found. Comput. Math. 14 (2014) 63–83.
- [6] H.H. Bauschke, D.R. Luke, H.M. Phan, X. Wang, Restricted normal cones and the method of alternating projections: theory, Set-Valued Var. Anal. 21 (2013) 431–473.
- [7] J.M. Borwein, A.S. Lewis, Convex analysis and nonlinear optimization: theory and examples, Springer, New York, 2006.

- [8] J.M. Borwein, D.R. Luke, Entropic regularization of the ℓ_0 function, in: H.H. Bauschke, R. Burachik, P.L. Combettes, V. Elser, D.R. Luke, H. Wolkowicz (Eds.), *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Springer, New York, 2011, pp. 65–92.
- [9] J.M. Borwein, M.K. Tam, Reflection methods for inverse problems with application to protein conformation determination, Springer volume on the CIMPA school *Generalized Nash Equilibrium Problems, Bilevel programming and MPEC* New Delhi, India, Dec. 2012., [arXiv:1408.4213](#).
- [10] E.J. Candes, K.R. Justin, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, 58 (8) (2006) 1207–1223.
- [11] A. Daniilidis, A. Lewis, J. Malick, H. Sendov, Prox-regularity of spectral functions and spectral sets, *J. Convex Anal.* 15 (3) (2008) 547–560.
- [12] D.L. Donoho, For most large underdetermined systems of linear equations the minimal ℓ_1 -norm solution is also the sparsest solution, *Comm. Pure Appl. Math.* 59 (2006) 797–829.
- [13] D. Drusvyatskiy, C. Kempton, Variational analysis of spectral functions simplified, [arXiv:1506.05170](#) (2015).
- [14] A. Eberhard, V. Roshchina, Orbital geometry in optimisation, [arXiv:1410.7940](#) (2015).
- [15] V. Elser, Phase retrieval by iterated projections, *JOSA A* 20 (1) (2003) 40–55.
- [16] J.R. Fienup, Phase retrieval algorithms: a comparison, *Appl. Opt.* 21 (1982) 2758–2769.
- [17] S. Foucart, H. Rauhut, *A Mathematical Introduction to Compressed Sensing*, Springer, New York, 2013.
- [18] T.L. Hayden, J. Wells, Approximation by matrices positive semidefinite on a subspace, *Lin. Alg. Appl.* 109 (1986) 115–130.
- [19] R. Hesse, D.R. Luke, P. Neumann, Alternating projections and Douglas–Rachford for sparse affine feasibility, *IEEE Trans. Sig. Proc.* 62 (18) (2014) 4868–4881.
- [20] J.-B. Hiriart-Urruty, J. Malick, A fresh variational-analysis look at the positive semidefinite matrices world, *J. Optim. Theory Appl.* 153 (2012) 551–577.
- [21] B. Gärtner, J. Matousek, *Approximation algorithms and semidefinite programming*, Springer-Verlag, Berlin Heidelberg, 2012.
- [22] A.S. Lewis, Nonsmooth analysis of eigenvalues, *Math. Program.* 84 (1) (1999) 1–24.
- [23] A.S. Lewis: The mathematics of eigenvalue optimization, *Math. Program., Ser. B* 97 (2003) 155–176.
- [24] A.S. Lewis, D.R. Luke, J. Malick, Local linear convergence for alternating and averaged nonconvex projections, *Found. Comput. Math.* 9 (2009) 485–513.
- [25] A.S. Lewis, J. Malick: Alternating projections on manifolds, *Math. Oper. Res.* 33 (1) (2008) 216–234.
- [26] D.R. Luke, Prox-regularity of rank constraint sets and implications for algorithms, *J. Math. Imaging Vision* 47 (3) (2013) 231–238.
- [27] B.S. Mordukhovich, *Variational analysis and generalized differentiation I*, Springer-Verlag, Berlin Heidelberg, 2006.
- [28] B.K. Natarajan, Sparse approximate solutions to linear systems, *SIAM J. Comput.* 24 (2) (1995) 227–234.
- [29] H.M. Phan, Linear convergence of the Douglas–Rachford method for two closed sets, *Optim.* 65 (2) (2016) 369–385.
- [30] R.A. Poliquin, R.T. Rockafellar, L. Thibault, Local differentiability of distance functions, *Trans. Amer. math. Soc.* 254 (200) 5231–5249.
- [31] H.-D. Qi, X. Yuan, Computing the nearest Euclidean distance matrix with low embedding dimensions, *Math. Program. Ser. A* 147 (2014) 351–389.

- [32] H.S. Sendov, Variational spectral analysis, PhD Thesis, University of Waterloo, 2006.
- [33] J.A. Tropp, Just relax: Convex programming methods for identifying sparse signals in noise, IEEE Trans. Inf. Theory 52 (3) (2006) 1030–1051.